DOCUMENT RESUME

ED 068 523 TM 001 884

AUTHOR	Pfeiffer, Mark G.; And Others
TITLE	Cross-Cultural Scaling Studies in the Development of Probabilistic Teaching Performance Criteria Anchored to Utility and Time Scales.
INSTITUTION	La Salle Coll., Philadelphia, Pa.
REPORT NO	PL-TR-5
PUB DATE	70
NOTE	49p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Classroom Techniques; *Cross Cultural Studies; Evaluation Criteria; Higher Education; *Models; Norms; *Performance Criteria; Post Testing; Predictive Measurement; *Rating Scales; Teaching Models; Training Techniques
IDENTIFIERS	Germany

ABSTRACT

Previous cross-cultural scaling studies are incorporated into a training and performance model of the classroom teaching job of the college professor. The model, based on German and American data, describes and sets a norm for the improvement of classroom performance. Eight teaching performance factors and a set of seven continua (six predictors and one criterion variable) for the quantification of teaching performance are included in the model. The present studies demonstrate the feasibility of normative teaching performance criteria anchored to physical scales. These scales may be applied to develop teacher training programs, and posttraining performance evaluation programs designed to improve classroom teaching. (Author)

ED 068523

PSYCHOLOGICAL LABORATORY
TECHNICAL REPORTS

NUMBER 5

1970

REVISED

TM 001 884

LA SALLE COLLEGE

PHILADELPHIA, PENNSYLVANIA

Cross-Cultural Scaling Studies in the Development of Probabilistic Teaching Performance Criteria Anchored to Utility and Time Scales[1]

Mark G. Pfeiffer

La Salle College, USA

Wolfgang Lehmann and Ulrich Scheidt

Ruhr University, W. Germany

Abstract

Previous cross-cultural scaling studies are incorporated into a training and performance model of the classroom teaching job of the college professor. The model, based on German and American data, describes and sets a norm for the improvement of classroom performance. Eight teaching performance factors and a set of seven continua (six predictors and one criterion variable) for the quantification of teaching performance are included in the model. The present studies have demonstrated the feasibility of normative teaching performance criteria anchored to physical scales. These scales may be applied to develop teacher training programs, and posttraining performance evaluation programs designed to improve classroom teaching.

---

[1]This is the fourth report in a series supported by an Alexander von Humboldt fellowship award to the first author, a Visiting Professor during the 1968-1969 academic year at the Ruhr University, West Germany. Slightly different versions of the second and third chapters of this report served as the theses of the junior authors for the Diploma in Psychology. The senior author conceived, designed and directed both of these thesis projects. Appreciation is due to Wolfhardt Matthaeus and Allan Nash for constructive hints, to Larry Rosbach for assistance in data collection and data analysis, and to Guenter Keim for graphics support. Data analyses were performed on an IBM 1620 computer (USA) and a TR-4 computer (Germany). Computational assistance was received from James Dougherty, Stanley Goliasczewski and Wolfgang Hawel. Reuben L. Hann proofed the final manuscript.

According to Gage (1963, Pp. 118-120) hundreds of studies have already been completed for the purpose of developing adequate criteria of teaching effectiveness. But these studies have largely yielded the disappointment of poor agreement from one study to the next and the general lack of psychological and educational meaning.

One solution to this criterion consistency problem could be the development of criteria of effectiveness in small, specifically defined aspects of the role of the teacher, rather than seeking criteria for the over-all effectiveness of teachers in the many varied facets of their roles (Gage, 1963).

The development of such specific aspects of teaching performance criteria are best exemplified in the studies which have succeeded in breaking up the rather complex whole of teaching behavior into highly specific and operationally defined classroom teaching factors, e.g., Isaacson et al. (1964); Pfeiffer and Rosbach, (1969).

Description of the Teaching Model

The approach to understanding the college professor's job adopted in the present report was to view it directly from the standpoint of the activities performed in the classroom. A multidimensional scaling analysis of classroom activities of psychology professors isolated these performance factors (Pfeiffer & Rosbach, 1969). Judgments were made by professors and by students of psychology. Thus, data were obtained from two different points of view, that of the student and that of his teacher. In both of these studies the same eight activity factors were found to account for the total classroom job. The high degree of congruence which was established between the factors developed independently from faculty and student

estimates opened the door for the further application of these eight factors.

Teaching Factors

Knowledge Dissemination

a. Distinguishes between fact and opinion
b. Presents application of theory
c. Employs textbook and/or prepared notes during lecture
d. Informs about information channels and sources (e.g., library)
e. Conducts research (e.g., collects data in class)

Teacher-Student Feedback

a. Provides feedback on tests and other material
b. Requests students to critique course
c. Conducts question-and-answer periods
d. Gives tests and quizzes
e. Responds to student questions

Advisory Guidance

a. Advises on vocational goals
b. Schedules student consultation

Information Dissemination

a. Gives handouts (e.g., course outline, etc.)
b. Writes on blackboard
c. Gives special instruction and information concerning labs, papers, etc.
d. Assigns outside readings and preparation (other than text)

Teacher Dynamism

a. Gesticulates and/or moves around while lecturing
b. Emphasizes material using humor
c. Gives examples from personal experience
d. Maintains eye contact

Control of Student Behavior

a. Responds to potential, or actual emergencies (e.g., student illness, bomb scare)
b. Administers school regulations (dress, smoking, etc.)
c. Transmits messages for others (e.g., departmental activities to be attended)

Classroom Administration

a. Assigns seating arrangement
b. Takes roll
c. Assigns grades
d. Establishes range of acceptable classroom behavior

Teaching Factors (cont'd)

Environmental Regulation

a. Checks physical equipment and environment (temperature, lights, etc.)
b. Operates equipment (e.g., audio-visual aids)
c. Accounts for school funds and property

It remained to quantify these dimensions along some set of continua so that the resultant representational model would possess a degree of precision not found within the more descriptive approaches (e.g., Siegel & Siegel, 1967). The integration of the activity factors, and the training correlates into a presentational format descriptive of the psychology professor's job is shown as Figure 1. Typically, models have been of three types: (1) the dimensional model, (2) the hierarchical model, and (3) the matrix model. Of these three, a combination of the hierarchical and matrix models seemed most appropriate as a presentational form.

The model as shown in Figure 1 quantitatively orders and organizes the classroom job of the psychology professor in that it shows the teaching activities ordered from low to high in terms of the criterion, Training Hours Required. Since each of the predictors (Training Correlates) was moderately to highly correlated with the criterion, this ranking of the job factors based on criterion data is fairly consistent for all variables investigated.

The model states that knowledge dissemination and information dissemination require the most training hours to achieve a norm of mastery and also that psychology professors spend most of their alloted classroom time with these types of activities. At the low end of this training time continuum are the factors, Classroom Administration and Environmental Regulation. The model also states that the predictors of Training Hours,

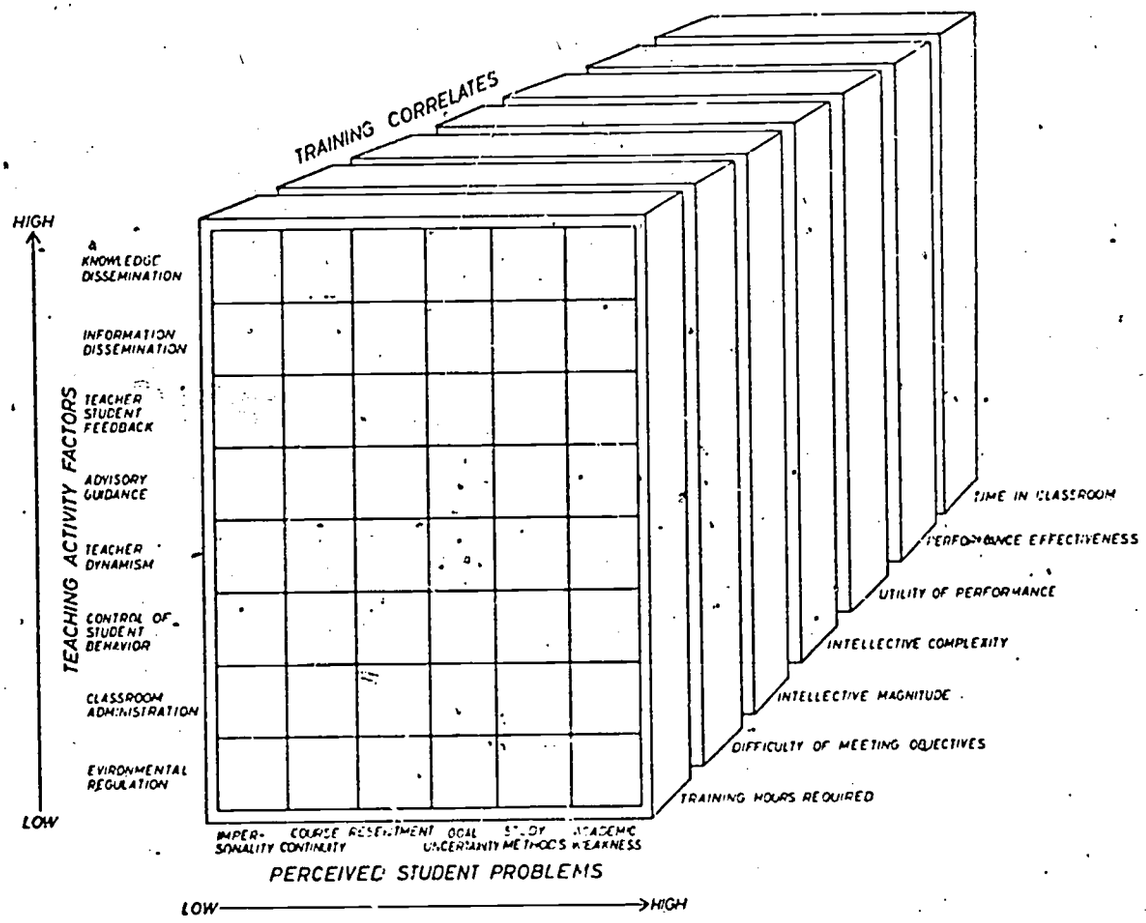

Figure 1: Structure of Teaching

(Student Problems are not discussed in this report, see Pfeiffer & Rosbach, 1968)

Difficulty of Meeting Objectives, Intellective Magnitude, Utility, etc. may be obtained via student or faculty magnitude estimations. These predictive validity coefficients have been in the .50 to .90 range.

Possibilities for the quantification of the model are presented by the Training Correlates dimension. A study by Pfeiffer (1969) investigated the relation between the criterion variable Training Hours Required and the following predictor variables: Difficulty of Meeting Objectives, Intellective Magnitude, Intellective Complexity, and Utility of Performance. These predictor variables were all quite highly correlated with the criterion (r=.78 to .88).

Training Correlates (quantified by students*)

The training correlates investigated for eight classroom jobs of psychology professors were:

1. Difficulty of meeting the teaching objective "readiness" -- a state of preparedness of teachers and teaching aids consistent with the requirements of the university, of accrediting agencies, and of present day society on a 0-100 scale of difficulty.

2. Utility of teaching performance -- how much money a semester of effective performance by a teacher is worth to the typical student for each of eight teaching jobs.

3. Intellective magnitude of teaching performance -- the average extent to which the 15 Guilford Structure-of-the-Intellect factors is involved in executing each of the eight classroom jobs of the teacher on a 0-100 scale.

4. Intellective complexity of teaching performance -- the number of Structure-of-the-Intellect cells at least moderately involved in executing each teaching job (maximum of 120 cells possible for each job).

* Except teaching time where teacher estimations were used in the final scale.

5. Performance effectiveness (e) -- the probability of effective performance (Pe) by a teacher based on the sum of effective performances divided by the sum of total performances (effective and ineffective) during one semester separately for each of eight teaching performance factors.

6. Teaching time required in the classroom -- the distribution of time spent by psychology professors in the performance of each of eight classroom teaching jobs. The sum of time estimations made by each S on each of the eight jobs had to equal 100 minutes.

Criterion Variable (quantified by teachers)

Number of training hours required to meet the objective "readiness" -- a state of preparedness of teachers and teaching aids consistent with the requirements of the school, of accrediting agencies, and of present day society for each of eight teaching jobs.

Preview of Further Studies

Chapter II reports the expected amount of teaching time wasted in the classroom and compares an American and a German psychology department. Estimations of teaching performance effectiveness were given by German and American psychology students on each of eight psychology teachers of their departments for each of eight classroom teaching performance factors. These values were converted to probabilities of ineffective teaching performance and multiplied with teaching time needed for each classroom teaching performance factor. By this procedure the teaching performance measurement of expected amount of time wasted in the classroom was developed.

Chapter III reports an investigation into the subjective expected utility of teaching performance. This investigation was also cross-cultural in that it compared American and German psychology departments.

Performance effectiveness estimations were given by German and American psychology students on each of eight psychology teachers of their departments for each of eight classroom teaching performance factors. These values were converted to probabilities and multiplicatively combined with the utility of teaching performance estimates of students to yield the subjective expected utility of classroom teaching performance. Intra- and inter-cultural similarities and differences are discussed.

CHAPTER II

Expected Amount of Teaching Time Wasted in the Classroom

Purpose

To achieve the purposes of the present study, the Structure-of-Teaching Model, shown as Figure 1 (Pfeiffer, 1969; Pfeiffer & Rosbach, 1969) was mathematically modified. Specifically the performance measure under development was the expected amount of time wasted in the classroom. Such an extension should enable the Structure-of-Teaching model to be used for normative as well as descriptive purposes.

It was hoped that, through the knowledge gained about the expected amount of time wasted in the classroom on each factor, one would know more exactly where to concentrate on improvement.

German and American departments were examined to get some idea as to how widely the performance measure under development might be generalized. It would seem that the expected amount of time wasted (or time used effectively) might be employed rather widely in making cross-cultural comparisons.

Method and Design

Stimuli

The following eight teaching performance factors, based on a multi-dimensional scaling analysis by Pfeiffer and Rosbach (1969), served as

stimuli: Knowledge Dissemination (KD), Information Dissemination (ID), Classroom Administration (CA), Environmental Regulation (ER), Advisory Guidance (AG), Teacher Dynamism (TD), Teacher-Student Feedback (TSF), and Control of Student Behavior (CSB).

Job Performance Measures

Time (t) expended by teachers. The time a teacher needs for performing each of the eight classroom teaching jobs was estimated by the method of constant sum, with the constraint that the sum of estimated time over all eight job factors equals 100 Minutes. With the assumption that these estimates represented a ratio scale of measurement, the values were directly convertable to percentages.

Teaching performance effectiveness (Pe). The effectiveness (e) a teacher achieves in his main subject during one semester was estimated by each S on a 0 to 100% scale with the constraint, that the effective (e) plus ineffective (i) performances equals 100 for each job factor. The teaching performance effectiveness measure was then defined by summing across Ss' judgments as follows:

$Pe = \sum e / \sum (e+i)$. Pe was then accepted as a probability value.

Expected amount of teaching time wasted (ETi). This measure was developed by combining time (t) multiplicatively with teaching performance ineffectiveness (Pi), where $Pi = 1 - Pe$. Thus $ETi = (t) \cdot (Pi)$. It should be noted that the Pi-and t-values were derived from independent subject groups.

Subjects

Four American and three German groups of Ss from an American and a German psychology department served as raters.

Americans. One of the American rater groups consisted of 12 American university teachers of psychology at a medium sized American city college in the Delaware Valley. This group made teaching time estimations (t). Additionally, a control group of 17 third and fourth year undergraduate psychology students from the same department made the same estimations on eight of these 12 teachers. Another independent group consisted of 17 third and fourth year undergraduate psychology students from the same city college. These Ss made performance effectiveness estimations (Pe) for eight of the teachers in their department. A third group of six American fourth year undergraduate psychology students served as controls by making the Pe-estimations by a different procedure.

Germans. Sixteen German university teachers of psychology at a medium sized university in the Ruhr area made the teaching time estimations (t). Additionally, a control group of 14 fourth and fifth year Vordiplom psychology students from the same department made the same estimations (t) on eight of these 16 teachers. A third group of Ss consisted of 17 German psychology students with essentially the same amount of psychology training as the other group of German students. They made separate performance effectiveness estimations (Pe) for each of eight teachers in their department.

Questionnaire Organization and Administration

The questionnaire format and the English language for all time estimations (t) and performance effectiveness estimations (Pe) was held constant across both cultures under investigation. No time limitation was given; however, none of the Ss needed more than 40 minutes to complete the questionnaire. Examples of the questionnaires are presented in Appendices A and B.

Time questionnaire. (See Appendix A). Time estimations (t) of the teaching time needed for performing each of the eight job factors were

given by American and German teachers concerning their own teaching activities and also by control groups of German and American students concerning the teaching activities of eight of their teachers.

In order to control effects due to presentation sequence and related variables, different questionnaire forms were developed. The two forms were assigned randomly to the teacher Ss. Each teacher made his estimations on one of two alternate questionnaire forms, containing opposite orders of the eight job stimuli.

Each student S made time estimations on one of four different questionnaire forms. These were assigned randomly to the Ss. The four forms differed in the ordering of the eight teachers to be judged. Additionaily the sequence of the given stimuli alternated from one page of the questionnaire to the next one.

Performance effectiveness questionnaire. (See Appendix B). The teaching performance effectiveness of eight teachers on each of the eight job factors was estimated by American and German student Ss.

Each S made estimations on one of four different questionnaire forms, containing four randomized orders of the teachers to be judged. Additionally the sequence of the given stimuli alternated from one page of the questionnaire to the next one. The four forms were assigned randomly to the Ss.

Results

Reliability of Scales

For achieving some idea of the stability of the scales, reliability coefficients (r_k) were computed by the ANOV procedures described in Winer (1962, p. 131, formula 4'). This reliability coefficient (r_k) is based on interjudge agreement and is interpretable as the most likely correlation

between the scale of the present sample and the scale to be derived from another sample of the same size drawn from the same population.

Time estimations (t). The reliability (interjudge agreement) of the time scale of 12 American teachers was $r_k = .97$, for the 16 German teachers $r_k = .97$. The German student Ss who served as a control group, estimated the performance of the sample of eight of their teachers. These reliabilities ranged from .80 to .98. Fisher's z' transformation was employed in computing the mean of these values ($\bar{r}_k = .96$). For the American students who served as a control group a more restricted range of reliabilities was found (r_k from .95 to .99; $\bar{r}_k = .98$).

The high agreement among the subjects within groups permitted the averaging of their data to yield four time scales, two for Americans and two for Germans (see Table 1).

Performance effectiveness estimations (Pe). Reliability estimates were also computed for the performance effectiveness scales (e) developed from the 17 American and 17 German students concerning eight of their teachers. For the eight American teachers rated by their students, the eight interjudge reliabilities ranged from .61 to .93 and for the German student rater group the eight reliabilities ranged from .76 to .91. The averaged reliabilities from these two rater groups based on Fisher's z' transformation were .79 and .86 for American and German rater groups, respectively.

Validity of Time Scales

To estimate the validity of the time scales, a sample of eight German teachers was selected. Each of these eight teachers, who estimated his own time distribution, was then compared with the average of 14 students' ratings of these same teachers. The eight correlation coefficients expressing the

degree of agreement across student-teacher derived time scales ranged from .44 to .99. These eight correlations averaged after Fisher's z' transformation, yielded a mean validity of $\bar{r}$ = .80. For the American data the same technique was employed (r ranged from .87 to .98; $\bar{r}$ = .96).

The time scale derived from the estimations of the eight German teachers, who were judged by the control group of 14 German students, showed high agreement (r = .99) with that one derived from time estimations of the other eight teachers not judged by the 14 students who served as controls. Thus the eight teachers selected for the validity study did not differ from the other group of eight not selected. A similar comparison was not made for the American data since there were only 12 teachers who made the time estimations and they could not be divided into equal groups of eight each.

Time Scales

Table 1 shows the teaching time scales based on the averaged estimations of 16 German teachers, 12 American teachers, and the control groups of 14 German students and 17 American students for eight job factors. With the exception of the time estimations by German teachers on the job factor, TSF, the scales derived from separate rater groups showed a great deal of similarity. The product-moment correlations among the scale values shown in Table 1 ranged from .77 to .99.

Table 1
Distribution of Teaching Time in the Classroom as Estimated by Various Subject Groups based on 100 Minute Periods

Teaching Activities	American Students N = 17	American Teachers N = 12	German Students N = 14	German Teachers N = 16
KD	45.84	54.25	42.91	35.19
ID	16.08	13.67	18.73	12.31
CA	4.65	2.04	3.72	1.53
ER	2.30	3.04	2.93	3.13
AG	5.39	1.92	5.59	2.59
TD	9.53	10.42	6.64	7.63
TSF	12.88	13.58	15.22	35.94
CSB	3.34	1.08	4.38	1.69
Sum	100.01	100.00	100.12	100.01

Teaching Performance Effectiveness (Pe)

To study the differences of the teaching performance estimations (Pe), as estimated by American and German students, a two dimensional analysis of variance with repeated measures on one dimension was performed (Winer, 1962, Pp. 302-312) using the data presented on Table 2.

Inspection of the summary table of the analysis of variance (Table 3) shows both the cultures and job factors as significant. The interaction between these factors was not significant. The American students tended to see their teachers as more effective than did German students. Differences across jobs were demonstrated and these differences tended to remain constant across the two cultures examined.

Expected Amount of Time Wasted in the Classroom (ETi)

Table 4 shows the matrix of expected amount of time wasted in the classroom (ETi) developed from American and German teachers' time estimations (t) and from American and German students' performance effectiveness estimations (Pe).

Each cell of the matrix contains an ETi-value, derived by multiplying each Pi = (1 - Pe) value with the mean value of time estimations (t) given by 12 American and 16 German teachers (Table 1).

To study the differences of the developed scales a two dimensional analysis of variance with repeated measures on one dimension and Newman-Keuls tests for differences between means were performed (Winer, 1962, Pp. 302-312). Inspection of the summary table of the ANOV (Table 5) shows the cultures and job activity factors to be significant. The interaction between the two variables was also significant ($\alpha = .01$) and is plotted in Figure 2. Figure 2 shows rather clearly that the significance of the interaction between the two variables is due to the job factor Teacher-Student Feedback (TSF). This finding was also

Table 2

Probability of Effective Teaching Performance as Estimated by Students Rating Their Teachers

		Teaching Activity							
		KD	TSF	ID	TD	ER	CA	AG	CSB
American Teachers	1	.765	.750	.741	.790	.559	.524	.589	.559
	2	.866	.746	.808	.681	.731	.697	.585	.724
	3	.718	.676	.749	.577	.559	.606	.609	.588
	4	.705	.715	.635	.746	.518	.574	.592	.635
	5	.741	.686	.726	.647	.609	.521	.759	.601
	6	.878	.814	.872	.828	.623	.789	.555	.752
	7	.738	.703	.679	.439	.553	.656	.551	.606
	8	.680	.624	.694	.583	.568	.571	.562	.653
Mean		.761	.714	.738	.661	.590	.617	.600	.640

		Teaching Activity							
		KD	TSF	ID	TD	ER	CA	AG	CSB
German Teachers	1	.715	.636	.682	.518	.403	.435	.515	.397
	2	.679	.538	.788	.450	.532	.424	.597	.538
	3	.732	.647	.744	.497	.365	.476	.529	.529
	4	.632	.635	.682	.503	.444	.385	.524	.588
	9	.738	.603	.759	.400	.397	.453	.485	.506
	10	.541	.321	.544	.721	.450	.382	.300	.471
	11	.556	.521	.621	.374	.435	.391	.588	.324
	16	.612	.532	.618	.512	.412	.421	.521	.400
Mean		.651	.554	.680	.497	.430	.421	.507	.469

[1]Note.--A control group of six American Ss generated a similar matrix of data by a counting of effective and ineffective performances procedure. The degree of agreement of these two American matrices was r = .63.

[2]Note.--The nature of the probability continuum is presented as Appendix D.

Table 3

Analysis of Variance of American and German Pe-Data

Source of variation	SS	df	MS	F
Between subjects	0.869	15		
A (cultures)	0.621	1	0.621	35.085*
Subjects within cultures	0.248	14	0.018	
Within subjects	1.259	112		
B (jobs)	0.695	7	0.099	19.451*
AB	0.061	7	0.087	1.706
B X subjects within cultures	0.503	98	0.005	
Total	2.128	127		

*Note.- 0.01 significance level.

Table 4

Expected Amount of Teaching Time Wasted in the Classroom (ETI) as Based on 100 Minute Periods

		Teaching Activity							
		KD	TSF	ID	TD	ER	CA	AG	CSB
	1	12.75	3.40	3.54	2.19	1.34	0.97	0.79	0.48
	2	7.27	3.45	2.62	3.32	0.82	0.62	0.80	0.30
American	3	15.30	4.40	3.43	4.41	1.34	0.81	0.75	0.45
	4	16.00	3.87	4.99	2.65	1.47	0.87	0.78	0.40
Teachers	5	14.05	4.27	3.75	3.68	1.19	0.98	0.46	0.43
	6	6.62	2.53	1.75	1.79	1.15	0.43	0.85	0.27
	7	14.21	4.03	4.39	5.84	1.36	0.70	0.86	0.43
	8	17.36	5.11	4.18	4.34	1.31	0.88	0.84	0.38
Mean		12.95	3.88	3.58	3.53	1.25	0.78	0.77	0.39

		Teaching Activity							
		KD	TSF	ID	TD	ER	CA	AG	CSB
	1	10.03	13.12	3.92	3.68	1.87	0.87	1.26	1.02
	2	11.30	16.60	2.51	4.19	1.46	0.88	1.05	0.78
German	3	9.43	12.69	3.15	3.84	1.98	0.80	1.22	0.80
	4	12.95	13.12	3.92	3.79	1.74	0.94	1.24	0.70
Teachers	9	9.22	14.27	2.97	4.58	1.88	0.84	1.34	0.83
	10	16.15	24.40	5.62	2.13	1.72	0.95	1.82	0.89
	11	15.62	17.21	4.67	4.77	1.77	0.93	1.07	1.14
	16	13.65	16.82	4.70	3.72	1.84	0.89	1.24	1.01
Mean		12.29	16.03	3.94	3.84	1.78	0.89	1.28	0.90

Note.--The data may be interpreted as percentages of total time.

Table 5

Analysis of Variance of American and German ETi-Data

Source of Variation	SS	df	MS	F
Between subjects	174.637	15		
A (cultures)	95.545	1	95.545	16.912*
Subjects within cultures	79.092	14	5.649	
Within subjects	3055.518	112		
B (jobs)	2331.747	7	333.107	146.215*
AB	500.505	7	71.501	31.385*
B X subjects within cultures	223.266	98	2.278	
Total	3230.155	127		

* Note.- 0.01 significance level

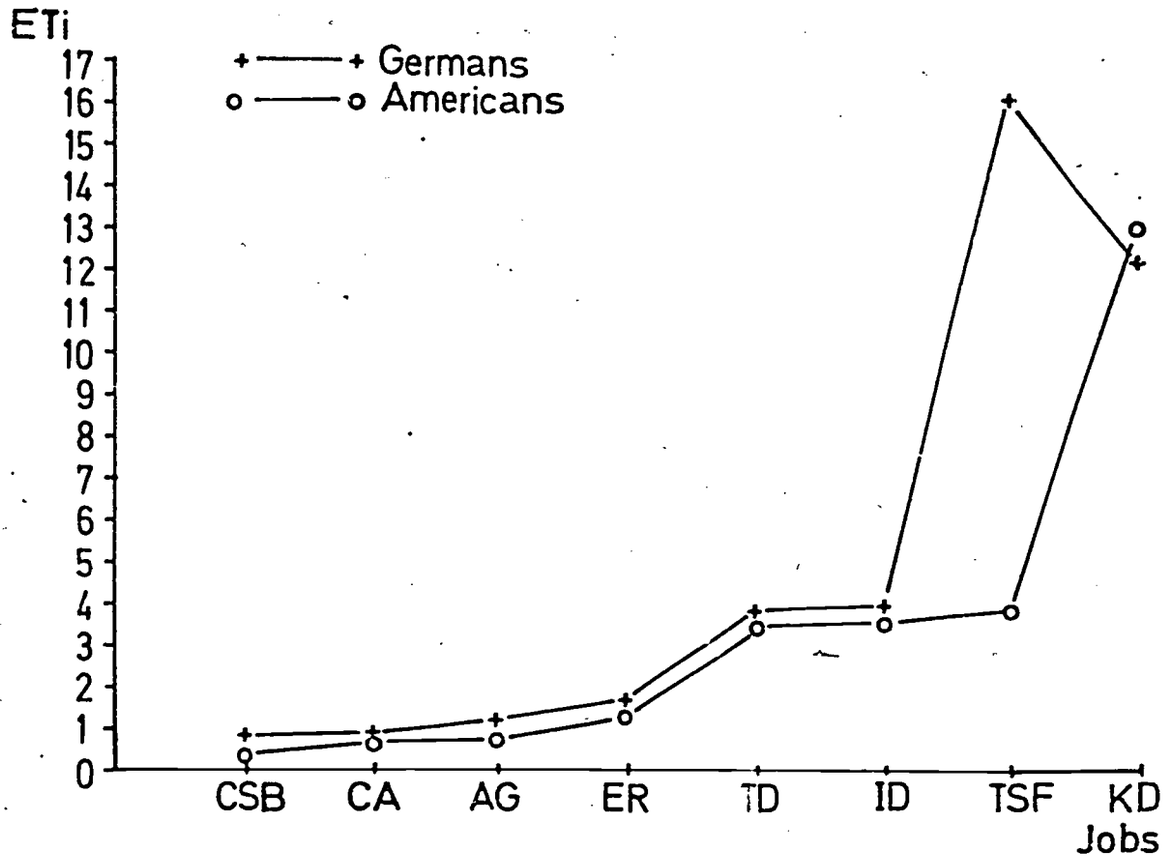

Figure 2. Expected percent of total teaching time wasted.

supproted by the results of Newman-Keuls tests. Despite the cross-cultural difference involving TSF, the product-moment correlation of the German-American ETi scales was moderately high ($r = .69$; $N = 8$ stimuli).

Additivity of the Pe-and-ETi Matrices

One method for determining the validity of the assumption of additivity (lack of interaction) which is inherent in matrix models is that of conjoint measurement. The theoretical basis of this technique has been developed by Luce and Tukey (1964). In the Guttman-Lingoes conjoint measurement 1 (CM-1) computer program (Lingoes, 1967) four theoretical matrices are developed from the input matrix. The models are developed via additive, subtractive, quadratic-additive and cubic-additive mathematical procedures. If the degree of correlation between the values in one of the theoretical matrices and the input matrix is high, it indicates that the mathematical model which holds for the development of the theoretical matrix also applies to the input matrix.

There were four matrices tested by these procedures. Each matrix was two dimensional and included eight teachers of a given department across one axis and the eight classroom jobs across the other axis ($N = 64$ cell entries).

Expected amount of teaching time wasted. CM-1 analyses were performed separately for the American and German ETi-data matrices. The theoretical vs. empirical correlations based on the additive model were, $r_{Americans} = .77$ and $r_{Germans} = .77$. These correlations were as high or higher than the outcomes of the other models tested.

Teaching performance effectiveness. The theoretical vs. empirical correlations based on the additive model were $r_{Americans} = .81$ and $r_{Germans} = .82$. These correlations were as high or higher than the outcomes of the other models tested.

Discussion

The reliability coefficients obtained for the collected data, suggested adequate stability of the Pe-scales. These findings agree with similar studies of technical skills by Siegel and Pfeiffer (1966) and Whitlock (1963).

The time scales also showed adequate reliability. For the time scales a validity estimate was also computed for both German and American samples. The magnitude of the averaged validity coefficients ($\bar{r}_{Americans} = .96$, $\bar{r}_{Germans} = .80$) compared favorably with a similar study involving the validity of time estimations by Carroll and Taylor (1969). Their averaged validity coefficients between estimated and actual time distribution of clerical workers' tasks was $\bar{r} = .88$. Carroll and Taylor reported that Stogdill and Shartle (1955) also found a fairly strong relationship between time estimates and actual logged time for specific work activities such as talking, reading, writing reports, and operating machines.

The results of the analysis of variance of ETi-data suggested that this performance effectiveness measure was useful in differentiating between psychology departments and job factors. The most interesting finding was the significant job-by-culture interaction. This interaction may be largely attributed to the cross-cultural difference in the expected amount of time wasted when engaging in Teacher-Student Feedback. The data suggest that the Germans wasted more time when executing this job than the Americans.

However, this finding was thrown open to question by the validity investigation of the German time distribution scale. Despite good overall student-teacher agreement in the scales, a large student-teacher disagreement existed for the factor Teacher-Student Feedback (Table 1). There were no student-teacher disagreements of this magnitude for the American time estimations.

It should be pointed out that an analysis of variance was also performed using the Pe-data alone. Here only the main effects were significant, i.e., the differences between the two cultures and among the eight teaching activity factors. Thus the significant interaction found in the ETi-data set involving the same independent variables can be attributed to the introduction of the time data alone and not the Pe-data.

The results of the conjoint measurement I analyses performed separately for German and American departments suggest that one can derive a measure of system effectiveness by summing across teachers and jobs to get a global view of how much time is wasted by the entire department. Thus with additivity of the matrix, ETi is both an individual as well as a system performance measure. The application of this principle as a normative performance measure is now clear. A single principle for the improvement of individual and department effectiveness exists, i.e., minimize the expected amount of teaching time wasted in the classroom and select factors where most time is wasted for possible improvement (Table 4). However, there may be a possibility that zero time wasted is in fact not optimal for a given system.

Conclusions

The results suggest the following conclusions:
(a) A model of teaching performance has been modified to include normative performance criteria, (b) The normative criteria are based on the expected amount of time wasted in the classroom by psychology teachers for each of eight jobs, (c) Adequate validity of the time scales is suggested by the high correlation between German and American

student-faculty time estimates. However validity of the composite performance effectiveness measure (ETi) has not been established, (d) Generalizability of the mathematical technique is suggested by a similar rank ordering of job effectiveness across German and American samples, (e) Because of additivity, the conjoint measurement analyses of the ETi-data suggest, at least for the continuum of Time Wasted in the classroom, how a system effectiveness measure for jobs performed may be derived by averaging across teachers within a given department. Naturally, the same statement would hold for the continuum of Time Used Effectively, (f) The normative performance measure developed may be useful for improving teaching effectiveness by showing where within a given educational system most teaching time is being wasted.

CHAPTER III

Expected Utility of Teaching Performance

Introduction

Following the views of Stevens (1959). there are two major conceptions which have dominated efforts in the field of utility theory. First, is a concept of utility which is called the classical view (see Stigler, 1950). This view was developed under the general assumption that money or wealth has a subjective value, and that with increasing dollars subjective value increases monotonically. The most widely accepted form of this function has been presented by Daniel Bernoulli (1738) who suggested that subjective value is a monotonically increasing and negatively accelerated function of the objective value. Thus the subjective value of money initially increases more rapidly than its objective value. In contrast to this concept of utility, which has no necessary relation to probability or risk, von Neumann and Morgenstern (1947) serve as representatives of a conception of utility built on a probabilistic view.

When calculating the expected utility (EU) the von Neumann and Morgenstern procedure implicitly assumes that the probabilities are known ahead of time and objectively determined (v. Neumann & Morgenstern, 1947). However more recently, and within a utility maximization model, subjective probabilities have been employed to derive subjective expected utilities (SEU). The SEU model (Edwards, Lindman & Phillips, 1965) multiplicatively combines subjective value (utility) with subjective probability to yield subjective expected utility.

Purpose

The general purpose of this study was to modify an aspect of the model shown as Figure 1 by multiplicatively combining two of the training correlates: Performance effectiveness and Utility of performance. By this procedure the potential of extending the descriptive model to a normative model, in the sense of showing the teacher how to maximize the utility of teaching, will be explored. The comparison of the data collected in Germany and the USA is intended to give some insight into the generalizability of the technique here developed. As a general conceptual framework SEU theory seemed most appropriate.

Method

Eight teaching activity factors based on a multidimensional scaling analysis (Pfeiffer & Rosbach, 1969) were used as stimuli. These stimuli included, Knowledge Dissemination (KD), Information Dissemination (ID), Teacher-Student Feedback (TSF), Advisory Guidance (AG), Teacher Dynamism (TD), Control of Student Behavior (CSB), Classroom Administration (CA), and Environmental Regulation (ER).

Job Performance Measures

Three measures were developed to assess the performance of university teachers of psychology.

Utility (U). Students estimated the money (in Dollars or Deutsche Mark) they were inclined to spend for a semester of effective teaching performance on each of eight classroom teaching activities. A magnitude estimation procedure was employed.

Probability of effective teaching performance (Pe). The basic data involved the effectiveness (e) and ineffectiveness (i) a teacher was thought by his students to have achieved on each of the eight teaching performance factors during one semester. Effectiveness estimations were made on a 0 - 100% scale with one constraint, i.e., that the estimations of the effective and ineffective performances within each job factor and for each ratee must sum to 100. The ratings were then summed across Ss. Probability of effective performance (Pe) was then defined by $Pe = \{e / \sum(e+i)$. These data are identical with the data presented in Chapter II (Table 2).

Subjective expected utility (SEU). The U-and Pe-values estimated by German and American psychology students were multiplicatively combined in order to get a measure of the subjective expected utility of each teacher on each of the eight teaching performance factors (Edwards, Lindman & Phillips, 1965).

Subjects

Five groups of Ss representing two nations served as raters. With the exception of the time estimation data generated by German and American students and some small differences in sample size, these were the same Ss reported earlier.

Americans. The two American rater groups were third and fourth year undergraduate psychology students studying at a medium sized eastern city college in the Delaware Valley, USA. One group of 17 Ss made the performance effectiveness ratings (Pe) and the same group (with one absent) of 16 Ss made the utility estimations (U). An independent control group of six senior psychology students made Pe-ratings by a slightly different counting procedure.

Germans. The three German rater groups were third, fourth, and fifth year Vordiplom psychology students studying at a medium sized university in the Ruhr area of Germany and a group of 14 of their teachers. One group of 17 student Ss made the performance effectiveness ratings (Pe), and another group of 14 student Ss made the utility estimations (U). The control group of 14 German teachers made utility estimations in the same manner as the students.

Questionnaire Organization and Administration

Questionnaires in the English language using a magnitude estimation procedure for the U-data and a constant sum procedure for the Pe-data were filled out by the American and German Ss. No time limitation was given, however the administration time was about 40 minutes.

Performance effectiveness questionnaire. The Ss rated each of eight teachers on each of eight jobs performed in the classroom. Two forms of the questionnaire were developed which presented reversed orders of the classroom teaching activities in order to control for sequence effects. Half of the Ss filled out the questionnaire with one order and half of the Ss with the reversed order. The presentation sequence of the eight teachers who were rated was varied randomly. Appendix B presents the performance effectiveness questionnaire.

Utility questionnaire. The Ss performed the utility estimations on the basis of each of eight teaching performance factors (Appendix C). The German Ss performed their ratings on the basis of Deutsche Mark values which were later transformed into Dollar values at the exchange rate of 4.00DM = $1.00. The two forms of the questionnaire, which presented reversed orders of the classroom teaching activities, were developed in order to control sequence effects. Each form was filled out by half of the Ss.

Results

Reliability of Scales

Since the SEU-data were calculated by combining U and Pe multiplicatively, reliability coefficients were computed on the basis of separate U- and Pe-values.

An estimate of the reliability of the probability of effective teaching performance estimations (interjudge agreement) was obtained by subject group for each of the American as well as for each of the German teachers. To achieve this goal the analysis of variance (Winer, 1962, p. 131) procedure was employed. Each analysis of variance yielded a statistic, r_k (Formula 4', p. 131) which is interpretable as the likely correlation between the scale values of the present sample of Ss and the average of a random sample of Ss of the same size drawn from the same population. Table 6 lists the obtained reliability values separately for each of the eight teachers rated in both cultures. Inspection of this table indicates that the range of the reliabilities for the 17 German Ss was .76 to .91. For the 17 American Ss participating, the reliability coefficients range from .61 to .93.

The same statistical procedure was employed to estimate interjudge reliability coefficients for the utility estimations. For the American students the reliability coefficient was $r_k = .94$, for the German students $r_k = .88$.

Table 6

Reliability (r_k) of Performance Effectiveness Ratings by Students of their Teachers.

American Teachers		German Teachers	
Teacher 6	.931	Teacher 10	.905
Teacher 7	.851	Teacher 9	.898
Teacher 1	.838	Teacher 3	.888
Teacher 5	.813	Teacher 1	.873
Teacher 2	.778	Teacher 2	.860
Teacher 4	.667	Teacher 11	.841
Teacher 3	.649	Teacher 4	.810
Teacher 8	.613	Teacher 16	.762

Note.- Reliability coefficients are ordered from high to low.

Validity of the Scales

Comparisons between the estimated money the student was inclined to pay for the effective performance of the eight classroom teaching activities and the real costs for a semester of study was consistent with the assumption that utility estimations were made on the basis of real-world values.

American students wished to pay 1221 Dollars (Table 7) for the effective performance of the eight teaching activities per semester. As listed in the 1967-1968 College Bulletin of their college the costs for a half year of study amounts to about 1100 Dollars. With 458 dollars per semester, the utility estimations of the German students ranged on a somewhat lower level. According to the German Hochschulführer of 1969, studying in Germany for one semester costs about 2000 Deutsche Mark (500 Dollars).

Finally, the degree of agreement of the utility scales independently generated by German students and teachers ($r = .89$) indicated an acceptable correspondence in the rank ordering of the jobs according to the utility of teaching (Table 7).

Additivity of the SEU Scales

In order to determine if teachers and teaching activities dimensions were statistically independent of each other, a Guttman-Lingoes (Lingoes, 1967) conjoint measurement analysis was performed using the data matrices presented as Tables 8 and 9. The Guttman-Lingoes program (CM-1) subjects the data to four tests to determine if: (1) the model is additive, (2) subtractive, (3) quadratic-additive, or (4) cubic-additive. CM-1 seeks to find a set of row and column values which, when combined in accordance with the four models, can be used to reproduce a new set of cell values which linearly correlate maximally with the original set of 64 cell values. The most parsimonious of the four mathematical models is the additive one (see Luce & Tukey, 1964).

Table 7

Means of Utility Estimations on Each Classroom Activity Factor

	Dollar Utility		
	American Students N = 16	German Students N = 14	German Teachers N = 14
Knowledge Dissemination	458	145	1124
Teacher-Student Feedback	218	84	191
Advisory Guidance	209	55	121
Information Dissemination	139	100	535
Teacher Dynamism	113	40	386
Environmental Regulation	39	16	42
Control of Student Behavior	27	10	12
Classroom Administration	18	8	26
Sum	1221	458	2437

Notes.--(1) American and German student data taken from Pfeiffer (1969).
(2) The product moment correlation coefficients expressing the degree of agreeement across these scales ranged from .85 to .91.

In terms of this measure of metric fit the additive model was as good or better than the other three models tested. The empirical vs. theoretical correlation of the scale values for the American data matrix was r = .97; for the German matrix r = .94. Thus the two dimensions could be interpreted as independent of each other (additive).

Additivity of the Performance Effectiveness Data

CM-1 was also performed for the American and the German Pe matrices and here again the results ($r_{Americans} = .81$; $r_{Germans} = .82$) showed that the dimensions are best fit to the model of additivity.

Analysis of Variance of the SEU Data

The combined data of Table 8 were treated variance analytically via a two-dimensional model with repeated measures on one dimension. This was done mainly in order to test the interaction between cultures and job activities and not the main effects. The significant difference between the American and the German departments is in part a function of the exchange rate between DM and $. The significance of the second main effect merely indicates that teaching activities can be differentiated.

An interaction of cultures and teaching activity factors can be observed (Table 9 ; Figure 3). The slope of the American data is steeper than the slope of the German data.

Table 8

Subjective Expected Utility

of Teaching Performance Per Semester

		Teaching Activity							
		KD	TSF	AG	ID	TD	ER	CSB	CA
	1	350.22	163.20	122.87	102.70	88.95	21.86	15.20	9.38
	2	396.45	162.33	122.03	111.99	76.68	28.58	19.69	12.48
American	3	328.70	147.10	127.04	103.81	64.97	21.86	15.99	10.85
	4	322.75	155.58	123.49	88.01	84.00	20.25	17.27	10.27
Teachers	5	339.23	149.27	158.33	100.62	72.85	23.31	16.35	9.33
	6	401.95	177.13	115.77	120.86	93.23	24.36	20.45	14.12
	7	337.86	152.97	114.94	94.11	49.43	21.62	16.48	11.74
	8	311.30	135.78	117.23	96.19	65.65	22.21	17.76	10.22
Mean		348.56	155.42	125.21	102.29	74.47	23.07	17.40	11.05

		Teaching Activity							
		KD	TSF	AG	ID	TD	ER	CSB	CA
	1	103.80	53.24	28.56	68.44	20.83	6.61	4.03	3.35
	2	98.58	45.11	33.10	79.08	18.10	8.72	5.46	3.26
German	3	106.27	54.24	29.33	74.67	19.99	5.98	5.37	3.66
	4	91.75	53.24	29.05	68.44	20.23	7.28	5.96	2.96
Teachers	9	107.14	50.56	26.89	76.17	16.09	6.51	5.13	3.49
	10	78.54	25.91	16.63	54.59	28.99	7.38	4.78	2.94
	11	80.72	43.68	32.60	62.32	15.04	7.13	3.29	3.01
	16	88.85	44.60	28.89	62.02	20.59	6.75	4.06	3.24
Mean		94.46	46.48	28.13	68.22	19.98	7.05	4.76	3.24

Note.--All cell entries are in U.S. Dollars.

Table 9

Analysis of Variance Summary of the

American and German SEU Data

Source of Variation	SS	df	MS	F
Between subjects	175691.89	15		
A (Cultures)	171214.89	1	171214.89	535.41*
Subjects within cultures	4477.00	14	319.78	
Within subjects	763326.06	112		
B (Jobs)	562329.87	7	80332.83	762.75*
AB	190673.91	7	27239.13	258.63*
B x subjects within cultures	10322.28	98	105.32	
Total	939017.95	127		

* Note.-o.ol significance level.

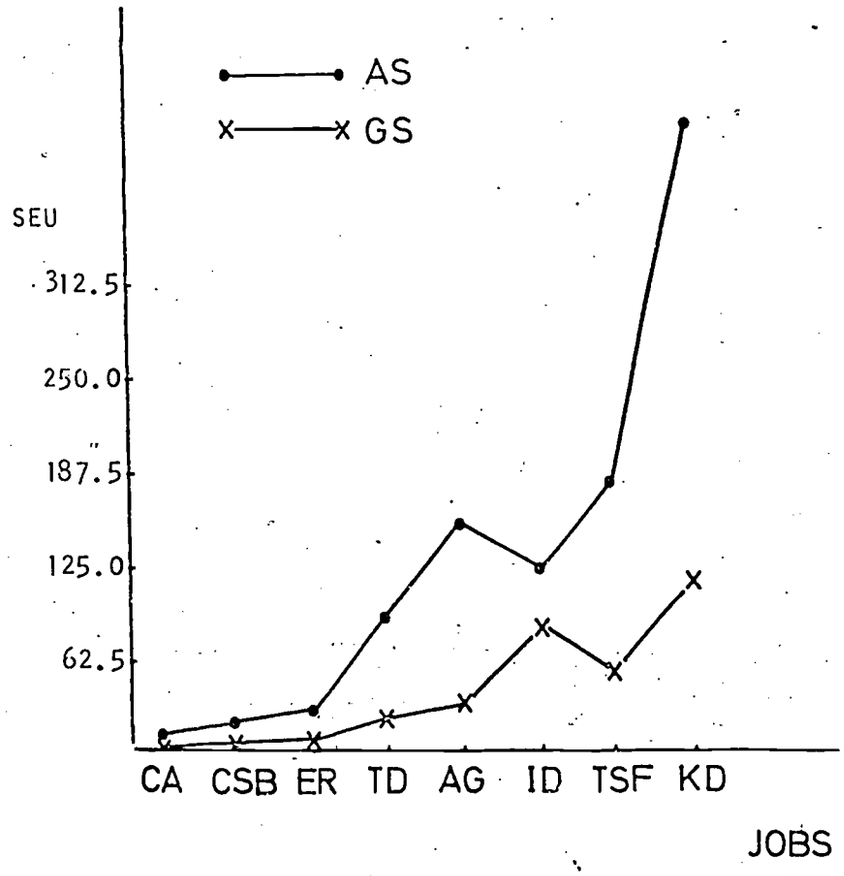

Figure 3. SEU of teaching performance for an average teacher during one semester.

Discussion

The list of interjudge reliability values suggests that there exists a moderate to very good stability in the Pe-as well as the U-estimations. Furthermore there was a moderate relationship between the magnitude of interjudge reliability and the effectiveness of the teachers on the SEU scale (r_A = .45; r_G = .52). Thus there was a tendency for greatest interjudge agreement where teaching effectiveness was also rated high.

The American and the German SEU scale values showed a similar rank order, with the American scale values at a higher level. The same non-parallel trend can be observed in Figure 4. Here again the averaged American U-estimations rise more steeply than the German estimations of teaching. The validation which has been performed for the student U-scale values gives reason for the statement that the cross-cultural differences in the U-estimations were partly due to the different levels of costs for one semester of study in the USA and Germany. USA students employed a range which was greater than the Germans, hence the curves had different slopes.

As can be examined in Table 9, the analysis of variance of the combined SEU-data shows significant main effects, and, which seems to be much more important, that an interaction existed between cultures and teaching activity factors. Inspection of Figure 3, which graphically presents these data, shows that there were some interesting cross-cultural differences in SEU-values. Advisory Guidance was valued more highly by American students than by German students. This cross-cultural difference held to an even greater extent for Knowledge Dissemination. Despite these differences there was a reasonably good cross-cultural correspondence in the rank order of the jobs within data sets.

An analysis of variance was also performed with the Pe-data alone. Here only the main effects were significant, i.e., the difference between

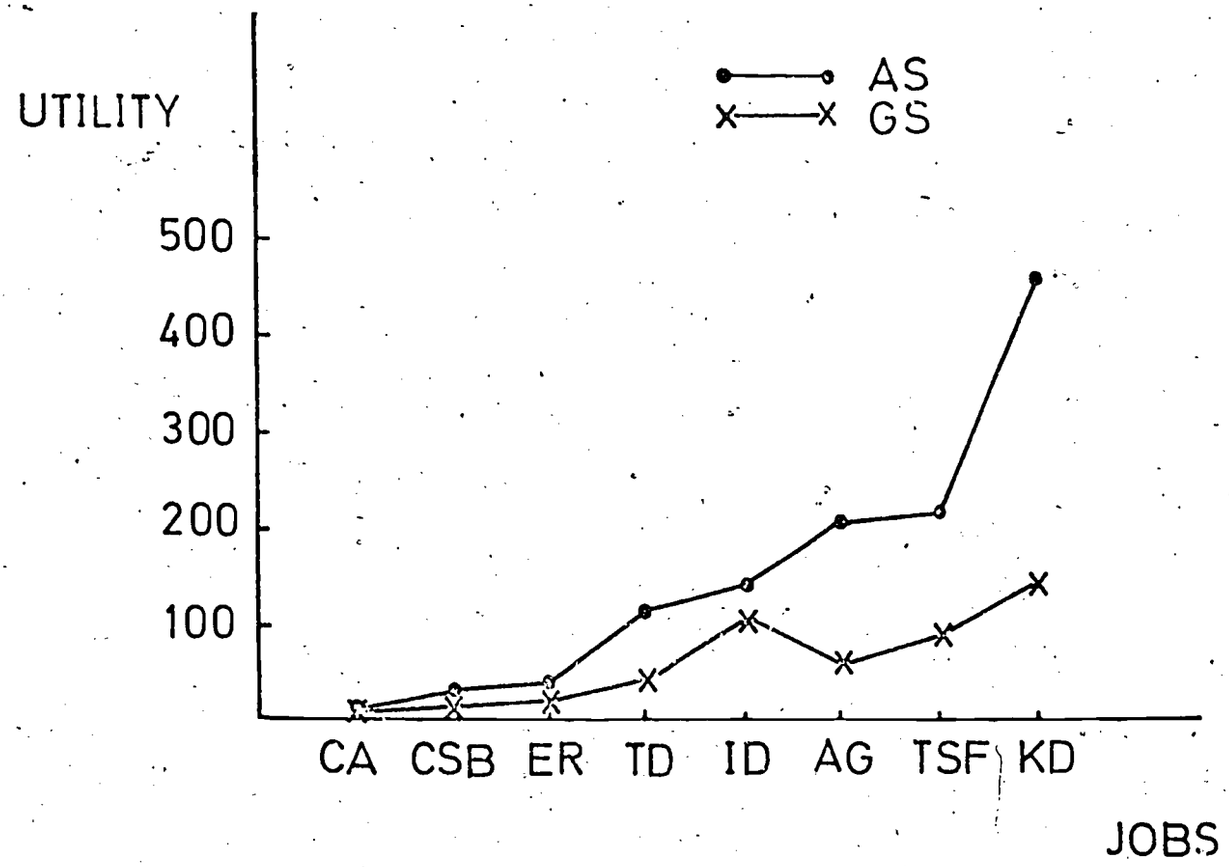

Figure 4. Dollar utility of effective teaching performance during one semester.

the two cultures and among the eight teaching activity factors. Thus the significant interaction found in the SEU-data set (Table 9) may be attributed to the introduction of the U-data and not the Pe-data. This conclusion is made obvious by a comparison of the Pe data (Figure 5) with the utility data (Figure 4). Figure 4 shows a non-parallel trend (based on a comparison of regression coefficients; Pfeiffer, 1969); however, the trend is parallel in Figure 5 (based on ANOV).

Additivity in the results of the conjoint measurement analyses performed separately for departments in this study suggests that a measure of system effectiveness can be derived by summing across teachers and jobs. Thus one could determine SEU-values of individual teachers across jobs, or of single jobs across teachers, or a single SEU-value for the entire department. When additivity exists, SEU is both an individual as well as a system performance measure. The single principle which exists for the improvement of individual and system effectiveness is the maximization of SEU. A practical upper limit could be established by determining the cost of improving the system, using the present values as a kind of base line.

From the practical standpoint, the presently achieved subjective expected utility measures allow the possibility of measuring the SEU of teaching in other science departments and to emphasize the teaching activities which are most highly estimated by students or faculty of that instructional system. Scaling analyses of this type could be done well in advance of the normal curriculum or instructional system development cycle (Pfeiffer, 1969; Pfeiffer & Siegel, 1966; 1967; Whitlock, 1963). Hence, one could properly anticipate areas for training emphasis.

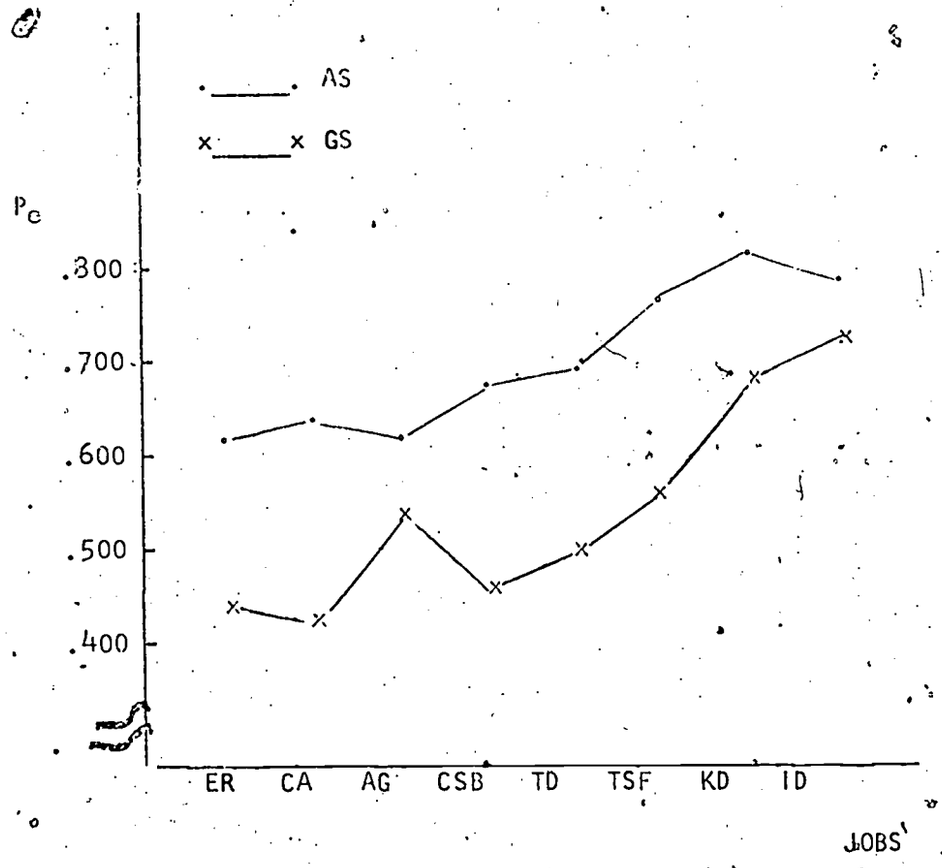

Figure 5. Pe-performance of an average teacher.

Conclusions

The results are interpreted to suggest the following conclusions: (a) a method has been developed to quantify the effectiveness of the classroom teaching job, (b) the SEU job performance measure achieved in this study was based on a multiplying procedure of students' utility (U) and probability of performance effectiveness (Pe) estimations, (c) the U-and Pe-estimations of the American as well as of the German raters showed a moderate to high interjudge reliability; utility estimations seemed to be sufficiently valid, (d) the SEU-scales may be employed to establish areas of training emphasis for psychology and other departments, (e) the generalizability of the technique is suggested by the fact that the SEU of teaching performance showed a similar rank ordering across German and American departments, and finally (f) the SEU extension of the teaching model may be employed as a norm to improve teaching effectiveness by showing the teacher exactly on which jobs improvement will yield maximum dividends consistent with the value system of students. Naturally the same procedures could be employed to establish a model consistent with the value system of teachers.

General Conclusions

The present studies have demonstrated that it is possible to develop normative teaching performance criteria that can be anchored to physical scales. In the present case these scales were time and money. The combination of these two scales may now be employed to define the optimal teacher, i.e., the one who is probably worth the most money and probably wastes the least amount of time. A similar definition could hold for the department, i.e., the group that probably wastes the least amount of time and is probably worth the most money. Thus the definition here developed contains elements of efficiency (performance over time) and value (money).

Future studies in this series will concentrate on developing the cost of improving the system i.e., by taking the present scale values as a base line one might determine the additional time and money necessary for improvement in a manner consistent with the subjective value systems of teachers and students. This technique would permit the establishment of an upper limit for improvement on each job factor.

Footnotes

[1]This is the fourth report in a series supported by an Alexander von Humboldt fellowship award to the first author, a Visiting Professor during the 1968-1969 academic year at the Ruhr University, West Germany. Slightly different versions of the second and third chapters of this report served as the theses of the junior authors for the Diploma in Psychology. The senior author conceived, designed and directed both of these thesis projects. Appreciation is due to Wolfhardt Matthaeus and Allan Nash for constructive hints, to Larry Rosbach for assistance in data collection and data analysis, and to Guenter Keim for graphics support. Data analyses were performed on an IBM 1620 computer (USA) and a TR-4 computer (Germany). Computational assistance was received from James Dougherty, Stanley Goliasczewski and Wolfgang Hawel. Reuben L. Hann proofed the final manuscript.

References

Bernoulli, D. Specimen theoriae novae de mensura sortis, Commentarii Academiae Scientiarum Imperialis Petropolitanae, 1738, 5, 175-192. (Transl. by L. Sommer, Econometrica, 1954, 22, 23-36).

Carroll, S.J. Jr., & Taylor, W.H. Jr. Validity of estimates by clerical personnel of job time proportions. Journal of Applied Psychology, 1969, 53, 164-166.

Edwards, W., Lindman, H. & Phillips, L. Emerging technologies for making decisions. In New directions in psychology, New York: Holt, Rinehart & Winston, 1965, 2, 259-325.

Gage, N.L. Paradigms for research on teaching. In Gage, N.L. (Ed.), Handbook of research on teaching, Chicago: Rand McNally & Company, 1963, Pp. 94-141.

Guilford, J.P. A study in psychodynamics. Psychometrika, 1939, 4, 1-23.

Isaacson, R.L., McKeachie, W.J., Milholland, J.E., Lin, Y.G., Hofeller, M., Baerwaldt, J.W., Zinn, K.L. Dimensions of student evaluations of teaching. Journal of Educational Psychology, 1964, 55, 344-351.

Lingoes, J.C. An IBM-7090 program for Guttman-Lingoes conjoint measurement 1. Behavioral Science, 1967, 12, 501-502.

Luce, R.D., & Tukey, J.W. Simultaneous conjoint measurement - a new type of fundamental measurement. Journal of Mathematical Psychology, 1964, 1, 1-27.

Neumann, J. v. & Morgenstern, O. Theory of games and economic behavior, Princeton: Princeton Univ. Press, 1947. (2nd ed.)

Pfeiffer, M.G. Training correlates of the psychology professor's job: agreement of scales across German and American psychology departments. Proceedings of the 77th Annual Convention of the American Psychological Association, 1969, 4. 643-644.

Pfeiffer, M.G., & Rosbach, L.A. Taxonomy of college student problems derived from a factor analytic interpretation of modified pair comparison scaled similarity judgments by students and their faculty. Philadelphia, Pa.: La Salle College, 1968. (Also in Proceedings of the 39th annual meeting of the Eastern Psychological Association, 1968).

Pfeiffer, M.G., & Rosbach, L.A. Teaching performance criterion development through scaling: Teacher-student points-of-view analysis. Perceptual and Motor Skills, 1969, 28, 755-766.

Pfeiffer, M.G., & Siegel, A.I. Personnel psychophysics, Perceptual and Motor Skills, 1966, 22, 257-258.

Pfeiffer, M.G., & Siegel, A.I. Personnel psychophysics in training systems. Proceedings of the 75th Annual Convention of the American Psychological Association, 1967, 2, 319-320.

Siegel, A.I., & Pfeiffer, M.G. Post training performance criterion development and application. Personnel psychophysics: estimating personnel subsystem reliability through magnitude estimation methods. Wayne, Pa.: Applied Psychological Services and Office of Naval Research Technical Report, 1966.

Siegel, L., & Siegel, L.C. A multivariate paradigm for educational research. Psychological Bulletin, 1967, 68, 306-326.

Stevens, S. S. Measurement, psychophysics, and utility. In Measurement: definitions and theories, Churchman, C. W. & Ratoosh, P. (Eds.), New York: Wiley, 1959, 18-63.

Stigler, G. J. The development of utility theory. Journal of Political Economy, 1950, 58, 307-327, 373-396.

Stogdill, R.M., & Shartle, C.L. Methods in the study of administrative leadership. Research Monograph 80, Bureau of Business Research, Ohio State University, Columbus, 1955.

Whitlock, G.H. Application of the psychophysical law to performance evaluation. Journal of Applied Psychology, 1963, 47, 15-23.

Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Appendix A

Sample Page of the Time Questionnaire

Time Distribution of Teaching Activities in the Classroom

Instructions

The purpose of this questionnaire is to determine how much time of a given total of 100 Minutes a psychology teacher spends on each of eight classroom teaching jobs.

A sheet of paper listing definitions of the classroom teaching activity factors is added.

There are no "right" and "wrong" answers. Your best answer is the only correct answer.

Remember that the sum of time over all classroom teaching activity factors must equal a total of 100 Minutes. Please confine your judgments to the time spent in the classroom only.

CLASSROOM TEACHING ACTIVITY	MINUTES
Knowledge Dissemination.....	_____
Information Dissemination.....	_____
Classroom Administration.....	_____
Environmental Regulation.....	_____
Advisory Guidance.....	_____
Teacher Dynamism.....	_____
Teacher-Student Feedback.....	_____
Control of Student Behavior.....	_____
	100 TOTAL

Appendix B

Sample Page of the Performance Effectiveness Questionnaire*

PERFORMANCE EFFECTIVENESS

The purpose of this part of the questionnaire is to determine the percent of the effective and ineffective performances in the classroom that you have observed among certain psychology teachers for one semester. Therefore your task will be to estimate the percent of effective and ineffective performances by a given teacher in his major subject for one semester. An example is provided to assist you in making these judgments.

Example

(Prof. John Doe)

CLASSROOM TEACHING ACTIVITY	PERCENT OF EFFECTIVE PERFORMANCES	PERCENT OF INEFFECTIVE PERFORMANCES
Application of Theory	75	25

Explanation

The person who completed the items felt that out of approximately 100% applications of theory in the classroom by Professor John Doe, 75 were effective and 25 ineffective.

Procedure

Enter your estimate of the percent of effective and ineffective performances, in dealing with his major subject, by _____ in the classroom over one semester for each activity listed. Refer to the definitions list for the meaning of each teaching activity.

CLASSROOM TEACHING ACTIVITY	PERCENT OF EFFECTIVE PERFORMANCES	PERCENT OF INEFFECTIVE PERFORMANCES
Knowledge Dissemination		
Information Dissemination		
Classroom Administration		
Environmental Regulation		
Advisory Guidance		
Teacher Dynamism		
Teacher-Student Feedback		
Control of Student Behavior		

* Adapted from Siegel & Pfeiffer (1966)

Appendix C

Sample Page of the Utility Questionnaire*

UTILITY

The purpose of this part of the questionnaire is to determine your opinion of the utility, or worth, of each teaching activity to the student. Your task is to estimate how much a semester of effective performance by a teacher in the classroom on each factor is worth to the typical psychology student. Enter your estimate of how much you would pay in dollars for a semester of effective performances on each factor. An example is given below to assist you in your determinations.

Example

Classroom Teaching Activities	Worth of One Semester's Effective Performance
Application of theory	$ 400.00
Taking roll	$ 10.00

Explanation

The person completing the questionnaire feels that a semester of effective performances by a teacher in "Application of theory" has substantial worth to the typical psychology student, giving it a value of $ 400.00. He feels that "Taking roll" is of much lower worth, giving it a value of $ 10.00.

Procedure

Enter your estimate of the worth, in dollars, to the typical psychology student, of effective performance by a teacher in the classroom of each factor. Refer to the definitions list for the meaning of each teaching activity. Remember that your estimate of worth should represent how much you would be willing to pay for one semester of effective performance.

CLASSROOM TEACHING ACTIVITY	WORTH OF ONE SEMESTER'S EFFECTIVE PERFORMANCE
Control of Student Behavior	$_____
Teacher-Student Feedback	$_____
Teacher Dynamism	$_____
Advisory Guidance	$_____
Environmental Regulation	$_____
Classroom Administration	$_____
Information Dissemination	$_____
Knowledge Dissemination	$_____

* Taken from Pfeiffer (1969)

Appendix D

Nature of the Probability Continuum